

# **Exhibit D**

## **Part 5**

# CMU's Incorrect Construction Covers Euclidean Metric

- “Euclidean branch metric” has noise samples that *vary together* [“identically” with “variance  $\sigma^2$ ”]

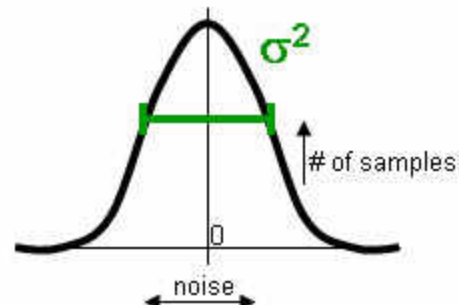
Euclidian branch metric. In the simplest case, the noise samples are realizations of **independent identically distributed** Gaussian random variables with zero mean and **variance  $\sigma^2$** . This is a white Gaussian noise assumption. This

'839 Patent 5:59-62

$$M_i = N_i^2 = (r_i - m_i)^2 \quad (8)$$

'839 Patent 6:12-13

Gaussian noise



- CMU argued to the Patent Office:

$$M_i = [r_i(0) - y_i(0)]^2 + [r_i(1) - y_i(1)]^2$$

**Such a branch metric is not correlation sensitive**, as claimed in independent claims 11, 16,

6/12/00 Amdt. at 9, '839 Patent File History  
(Marvell Exh. 22)

# CMU's "Disclosed Embodiment" Argument Fails

- CMU concedes: "Figure 3A calculates a correlation ... in one disclosed embodiment"

**FIG. 3A** illustrates a block diagram of a branch metric computation circuit **48** that **computes the metric  $M_i$  for a branch of a trellis, as in Equation (13).** Each branch of the

CMU Reply at 4

'839 Patent 7:14-18

- Figure 3A computes Equation (13)
- Using Noise Covariance Matrix  $\hat{C}$
- The  $\hat{C}(\hat{a})$  estimate calculates correlation:  
$$E[\hat{C}(\hat{a})] = E[N_i N_i^T]$$
- Marvell covers this embodiment**

need for further mean corrections. The focus is shifted to tracking the noise covariance matrices needed in the computation of the branch metrics (13).

Assume that the sequence of samples  $r_i, r_{i+1}, \dots, r_{i+L}$  is observed. Based on these and all other neighboring samples, after an appropriate delay of the Viterbi trellis, a decision is made that the most likely estimate for the sequence of symbols  $a_{i-K_f}, \dots, a_{i+L+K_f}$  is  $\hat{a}_{i-K_f}, \dots, \hat{a}_{i+L+K_f}$ . Here  $L$  is the noise correlation length and  $K=K_f+K_r+1$  is the ISI length. Let the current estimate for the  $(L+1) \times (L+1)$  covariance matrix corresponding to the sequence of symbols  $\hat{a}_{i-K_f}, \dots, \hat{a}_{i+L+K_f}$  be  $\hat{C}(\hat{a}_{i-K_f}, \dots, \hat{a}_{i+L+K_f})$ . This symbol is abbreviated with the shorter notation,  $\hat{C}(\hat{a})$ . If the estimate is unbiased, the expected value of the estimate is:

$$E\hat{C}(\hat{a}) = E[N_i N_i^T] \quad (21)$$

where  $N_i$  is the vector of differences between the observed samples and their expected values, as defined in (12).

'839 Patent 9:21-37

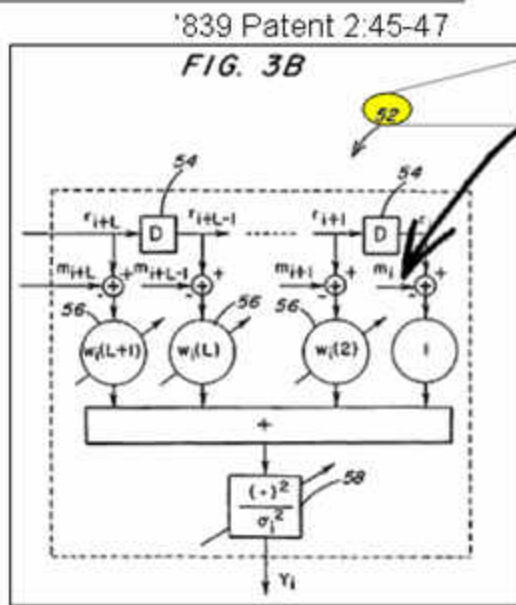
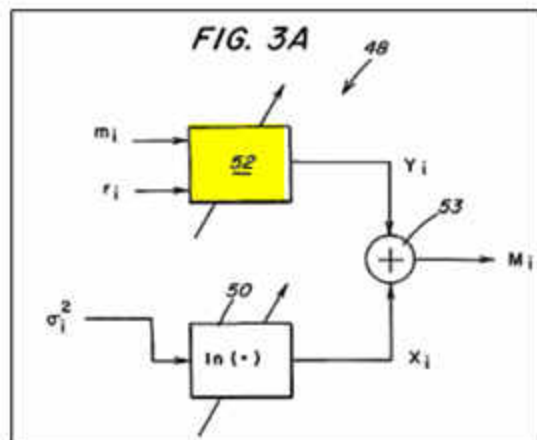
# CMU's "Disclosed Embodiment" Argument Fails

- McLaughlin: "an embodiment of which is shown in Figure 3B of the CMU Patents ... without using or computing the expected value of the product of the signal samples."

McLaughlin Decl. at ¶ 13

- Fig. 3B only shows one implementation of Circuit 52 in Fig. 3A

FIG. 3B is an illustration of an implementation of a portion of the branch metric computation module of FIG. 3A;





Claim Term	CMU's Construction	Marvell's Construction
<p>covariance</p> <p>'839 Patent Claims 11, 16, 19, 23 '180 Patent Claim 6</p>	<p>none; see "noise covariance matrices"</p> <p>CMU Brf. at 27</p>	<p>the expected (mean) value of the product of <math>(r_i - m_i)</math> and <math>(r_j - m_j)</math>, where <math>r_i</math> and <math>r_j</math> are observed signal samples (at time <math>i</math> and time <math>j</math>, respectively) and <math>m_i</math> and <math>m_j</math> are the expected (mean) values of the samples (at time <math>i</math> and time <math>j</math>, respectively) (i.e., <math>E[(r_i - m_i)(r_j - m_j)]</math>).</p> <p>Marvell Brf. at 21-24</p>

- Dispute
  - Should "covariance" have its ordinary meaning (Marvell) or be read out of the claim (CMU)?

delayed samples and detector decisions to update the noise statistics, i.e., to update the noise covariance matrices. A metric computation update circuit 36 uses the updated statistics to calculate the branch metrics needed in the Viterbi-like algorithm. The algorithm does not require replacing current detectors. It simply adds two new blocks

for a feedback circuit 32 which feeds back into a Viterbi-like detector 30. The outputs of the detector 30 are decisions and delayed signal samples, which are used by the feedback circuit 32. A noise statistics tracker circuit 34 uses the delayed samples and detector decisions to update the noise statistics, i.e., to update the noise covariance matrices. A metric computation update circuit 36 uses the updated statistics to calculate the branch metrics needed in the Viterbi-like algorithm. The algorithm does not require replacing current detectors. It simply adds two new blocks to the feedback loop to adaptively estimate the branch metrics used in the Viterbi-like detector 30.

The Viterbi-like detector 30 typically has a delay associated with it. Until the detector circuit 30 is initialized, signals of known values may be input and delayed signals are not output until the detector circuit 30 is initialized. In other types of detectors, the detector may be initialized by having the necessary values set.

The correlation-sensitive maximum likelihood sequence detector (CS-MLS) 26 is described hereinafter. Assume that  $N$  channel bits (symbols),  $x_1, x_2, \dots, x_N$ , are written on a magnetic medium. The symbols  $x_i$ ,  $i=1, \dots, N$ , are drawn from an alphabet of four symbols,  $x_i \in \{+1, 0, -1, \psi\}$ . The symbols  $+$  and  $-$  denote a positive and a negative transition, respectively. The symbol  $\psi$  denotes a writer zero (no transition) whose nearest preceding non-zero symbol is a  $+$  while  $\psi$  denotes a writer zero whose nearest preceding transition is a negative one, i.e.,  $-$ . This notation is used because a simple transition of transitions as  $+-$  and no transitions as  $++$  is blind to signal asymmetries (MR head asymmetries and base line drifts), which is inappropriate for the present problem. In FIG. 3 a sample transition is illustrated. The signal asymmetries and base line shifts are exaggerated in FIG. 3. FIG. 3 also shows the writer symbols  $x_1, \dots, x_N$  as well as the samples  $r_1, \dots, r_N$  of

the received signal. The likelihood function (conditional pdf) is factored into a product of conditional pdfs:

$$p(r_1, \dots, r_N | x_1, \dots, x_N) = \prod_{i=1}^N p(r_i | x_1, \dots, x_N, r_{i-1}, \dots, r_1) \quad (2)$$

To proceed and obtain more concrete results, the nature of the noise and of the intersymbol interference is suggestive recording is exploited.

Finite correlation length. The conditional pdfs in Equation (2) are assumed to be independent of former samples after some length  $L$ , i.e.,  $L$  is the correlation length of the noise. This independence leads to:

$$p(r_1, \dots, r_N | x_1, \dots, x_N) = p(r_1 | x_1) p(r_2 | x_1, x_2) \dots p(r_N | x_{N-L+1}, \dots, x_N) \quad (3)$$

Finite intersymbol interference. The conditional pdf is assumed to be independent of symbols that are not in the  $K$ -neighborhood of  $x_{i-L}, \dots, x_{i-L+K}$ . The value of  $K$  is determined by the length of the intersymbol interference (ISI). For example, for PR4,  $K=2$ , while for EPR4,  $K=3$ .  $K_L$  is defined as the length of the leading (transient) ISI and  $K_T$  is defined as the length of the trailing (transient) ISI, such that  $K=K_L+K_T$ . With this notation the conditional pdf in (3) can be written as:

$$p(r_1, \dots, r_N | x_1, \dots, x_N) = p(r_1 | x_1) p(r_2 | x_1, x_2) \dots p(r_N | x_{N-L+K_L}, \dots, x_N) \quad (4)$$

Substituting (4) into (2) and applying Bayes rule, the factored form of the likelihood function (conditional pdf) is obtained:

# Document 108-13 File



This correlation value is hard to place in context, so there is a related statistic, called covariance, which measures both the degree and the direction of the relationship between two sets of data. See Proakis Decl. at ¶ 22; P. Olofsson, Probability, Statistics and Stochastic

covariance, which measures both the degree and the direction of the relationship between two sets of data. See Proakis Decl. at ¶ 22; P. Olofsson, Probability, Statistics and Stochastic Processes, at 256-261 (2005) (Feb. 19). The covariance is zero if two sets of data are not related, positive if large data values in one set correspond to larger values in the other set, and negative if larger data values in one set correspond to smaller data values in the other set. Olofsson at 261. The "strength" the dependence between two sets of data, the larger the covariance. 56. As its name suggests, covariance is a combination of correlation and variance. Covariance is calculated by subtracting the mean for each data point to obtain deviations before the multiplication and then taking the average of the products. Pocket Dictionary of Statistics ("covariance") (John Wiley & Sons (Feb. 17 at 119). In the last example, the mean and deviations for Test 1, were calculated above. The mean of Test 2 is 80, and the deviations for both Test #1 and Test #2 are shown in Table 2 below. To calculate the covariance of these two

- “Covariance” is used independently without “noise covariance matrices”



10. A method of generating a branch weight for branches of a trellis for a Viterbi-like detector, wherein the detector is used in a system having Gaussian noise, comprising:

selecting a plurality of signal samples, wherein each sample corresponds to a different sampling time instant;

calculating a first value representing a logarithm of a quotient of a determinant of a trellis branch dependent covariance matrix of said signal samples and a determinant of a trellis branch dependent covariance matrix of a subset of said signal samples;

calculating a second value representing a quadratic of said signal samples less a plurality of target values normalized by a trellis branch dependent covariance of said signal samples;

calculating a third value representing a quadratic of a subset of said signal samples less a plurality of channel target values normalized by a trellis branch dependent covariance of said subset of signal samples;

calculating the branch weight from said first, second, and third values; and

'839 Patent at Claim 10



- Covariance Matrices used to calculate correlation-sensitive branch metrics

15.

computing said branch weight.

11. A method for detecting a sequence that exploits the correlation between adjacent signal samples for adaptively detecting a sequence of symbols through a communications channel having intersymbol interference, comprising the steps of:

(a) performing a Viterbi-like sequence detection on a plurality of signal samples using a plurality of correlation-sensitive branch metrics;

(b) outputting a delayed decision on the transmitted symbol;

(c) adaptively updating a plurality of noise covariance matrices in response to said delayed signal samples and said delayed decisions;

(d) recalculating said plurality of correlation-sensitive branch metrics from said noise covariance matrices using subsequent signal samples; and

(e) repeating steps (a)–(d) for every new signal sample.

12. The method of claim 11 wherein said sequence detection is performed using a PRM.

13. The method of claim 11 wherein said sequence detection is performed using an FDT.

14. The method of claim 11 wherein said sequence detection is performed using an RAM.

15. The method of claim 11 wherein said sequence detection is performed using an MDP.

16. A method for detecting a sequence that exploits the correlation between adjacent signal samples for adaptively detecting a sequence of symbols through a communications channel having intersymbol interference, comprising the steps of:

(a) performing a Viterbi-like sequence detection on a plurality of signal samples using a plurality of correlation-sensitive branch metrics;

(b) outputting a delayed decision on the transmitted symbol;

(c) outputting a delayed signal sample;

(d) adaptively updating a plurality of noise covariance matrices in response to said delayed signal samples and said delayed decisions;

(e) recalculating said plurality of correlation-sensitive branch metrics from said noise covariance matrices using subsequent signal samples; and

(f) repeating steps (a)–(e) for every new signal sample.

17. The method of claim 16 wherein said sequence detection is performed using a PRM.

18. The method of claim 16 wherein said sequence detection is performed using an FDT.

19. The method of claim 16 wherein said sequence detection is performed using an RAM.

20. The method of claim 16 wherein said sequence detection is performed using an MDP.

21. A detector circuit for detecting a plurality of data from said signals, said detector circuit having a circuit for

16. A method for detecting a sequence that exploits the correlation between adjacent signal samples for adaptively detecting a sequence of symbols through a communications channel having intersymbol interference, comprising the steps of:

(a) performing a Viterbi-like sequence detection on a plurality of signal samples using a plurality of correlation sensitive branch metrics;

(b) outputting a delayed decision on the transmitted symbol;

(c) outputting a delayed signal sample;

(d) adaptively updating a plurality of noise **covariance matrices** in response to said delayed signal samples and said delayed decisions;

(e) recalculating said plurality of correlation-sensitive branch metrics from said noise **covariance matrices** using subsequent signal samples; and

(f) repeating steps (a)–(e) for every new signal sample.

See '839 Patent Claims 11, 16, 19, 23; '180 Patent Claim 6

delayed samples and detector decisions to update the noise statistics, i.e., to update the noise covariance matrices. A metric computation update circuit 36 uses the updated statistics to calculate the branch metrics needed in the Viterbi-like algorithm. The algorithm does not require replacing current detectors. It simply adds two new blocks

for a feedback circuit 32 which feeds back into a Viterbi-like detector 30. The outputs of the detector 30 are decisions and delayed signal samples, which are used by the feedback circuit 32. A noise statistics tracker circuit 34 uses the delayed samples and detector decisions to update the noise statistics, i.e., to update the noise covariance matrices. A metric computation update circuit 36 uses the updated statistics to calculate the branch metrics needed in the Viterbi-like algorithm. The algorithm does not require replacing current detectors. It simply adds two new blocks to the feedback loop to adaptively estimate the branch metrics used in the Viterbi-like detector 30.

The Viterbi-like detector 30 typically has a delay associated with it. Until the detector circuit 30 is initialized, signals of known values may be input and delayed signals are not output until the detector circuit 30 is initialized. In other types of detectors, the detector may be initialized by having the necessary values set.

The correlation-sensitive maximum likelihood sequence detector (CS-MLS) 26 is described hereinafter. Assume that  $N$  channel bits (symbols),  $b_1, b_2, \dots, b_N$ , are written on a magnetic medium. The symbols  $b_1, b_2, \dots, b_N$  are drawn from an alphabet of four symbols,  $\{+1, 0, -1, \emptyset\}$ . The symbols  $+$  and  $-$  denote a positive and a negative transition, respectively. The symbol  $\emptyset$  denotes a writer zero (no transition) whose nearest preceding non-zero symbol is a  $+$  while  $\emptyset$  denotes a writer zero whose nearest preceding transition is a negative one, i.e.,  $-$ . This notation is used because a simple transition of transition  $+$  to  $-$  and no transitions in  $\emptyset$ 's is blind to signal asymmetries (MR head asymmetries and base line drifts), which is inappropriate for the present problem. In FIG. 3, a sample waveform is illustrated. The signal asymmetries and base line shifts are exaggerated in FIG. 3. FIG. 3 also shows the writer symbols  $b_1, \dots, b_N$  as well as the samples  $r_1, \dots, r_N$  of

the received signal. The likelihood function (conditional pdf) is factored into a product of conditional pdfs:

$$p(r_1, \dots, r_N | b_1, \dots, b_N) = \prod_{i=1}^N p(r_i | b_1, \dots, b_N, r_{i-1}, \dots, r_1) \quad (2)$$

To proceed and obtain more concrete results, the nature of the noise and of the intersymbol interference is suggestive recording is exploited.

Finite correlation length. The conditional pdfs in Equation (2) are assumed to be independent of former samples after some length  $L$ . I.e.,  $L$  is the correlation length of the noise. This independence leads to:

$$p(r_1, \dots, r_N | b_1, \dots, b_N) = p(r_1 | b_1, \dots, b_L, r_0, \dots, r_{-L}) \times p(r_2 | b_2, \dots, b_{L+1}, r_1, \dots, r_{-L+1}) \times \dots \times p(r_N | b_N, \dots, b_{N-L+1}, r_{N-1}, \dots, r_{N-L}) \quad (3)$$

Finite intersymbol interference. The conditional pdf is assumed to be independent of symbols that are not in the  $K$ -neighborhood of  $t_1, \dots, t_N$ . The value of  $K$  is determined by the length of the intersymbol interference (ISI). For example, for PR4,  $K=2$ , while for EPR4,  $K=3$ .  $K_c, 2D$  is defined as the length of the leading (trailing) ISI and  $K_s, 2D$  is defined as the length of the trailing (leading) ISI, such that  $K=K_c+K_s+1$ . With this notation the conditional pdf in (3) can be written as:

$$p(r_1, \dots, r_N | b_1, \dots, b_N) = p(r_1 | b_{1-K_c}, \dots, b_{1-K_s}, r_0, \dots, r_{-L}) \times p(r_2 | b_2, \dots, b_{L+1}, r_1, \dots, r_{-L+1}) \times \dots \times p(r_N | b_N, \dots, b_{N-L+1}, r_{N-1}, \dots, r_{N-L}) \quad (4)$$

Substituting (4) into (2) and applying Bayes rule, the factored form of the likelihood function (conditional pdf) is obtained:

# Claim Language: Independent Meaning

- “Covariance matrix” is used independently without “noise covariance matrices”

13  
US 6,201,000

shows the performance of the PBA detector at this density. FIG. 9 is similar to FIG. 7, except that the error rate has increased. This is again due to a mismatch between the original signal and the PBA target response, which is why the PBA shaping filter introduces distortion in the noise. PBA (2) will compensate the noise after adaptation, allowing the value of exploiting the correlation across signal samples.

FIG. 10 shows the error rate obtained when using the LPA detector. Due to a higher density, the error rate is higher than in the previous example with a target separation of 4.4. This is why the gain in FIG. 10 has moved to the right by 3 dB in comparison to the graph in FIG. 8. While the required SNR (dB) increased, the margin between the LPA (1) and LPA (2) also increased from about 0.5 dB to about 1 dB, suggesting that the correlation across signal samples is more robust to density increase. This is illustrated in FIG. 11 where the SNR (dB) required for an error rate of  $10^{-3}$  is plotted versus the target density for the three LPA detectors. From FIG. 11 it can be seen that, for example, with an SNR (dB) of 15, the LPA (1) detector operates at a linear density of about 2.2 symbols/PW30 and the LPA (2) detector operates at 2.4 symbols/PW30, thus achieving a gain of about 10% of linear density.

Symbol separation of 2.5. This recording density corresponds to a symbol density of 3 symbols/PW30. This is a very low number of symbols per s.d. as the density where the detector significantly loses performance due to the penetration of impulsive distortion, also referred to as nonlinear amplitude loss or partial signal noise. FIGS. 12 and 13 show the performance of the PBA and LPA families of detectors at this density. The detector with the L2 metric outperforms the other two metrics. The error rates are quite high in all cases. This is because at the symbol separation of 2.5, nonlinear effects, such as partial noise due to penetration of distortion, are in evidence. These effects can only be undone with a nonlinear pulse shaping filter, which has not been employed here.

The experimental evidence shows that the correlation across signal samples improves the performance of the correlation insensitive detector. It has also been demonstrated that the performance margin between the correlation sensitive and the correlation insensitive detector grows with the recording density. In other words, the performance of the correlation insensitive detector deteriorates faster than the performance of the correlation sensitive detector. Quantitatively, this margin depends on the amount of correlation in the noise passed through the system. Qualitatively, the higher the correlation between the noise samples, the greater will be the margin between the CS-SD and the correlation insensitive detector.

While the present invention has been described in connection with preferred embodiments thereof, many modifications and variations will be apparent to those of ordinary skill in the art. For example, the present invention may be used to detect a sequence that exploits the correlation between adjacent signal samples for adaptively detecting a sequence of symbols through a communication channel. The foregoing description and the following claims are intended to cover all such modifications and variations.

What is claimed is:

1. A method of determining branch metric values for branches of a trellis for a Viterbi-like detector, comprising: selecting a branch metric function for each of the branches at a certain time index; and applying each of said selected functions to a plurality of signal samples to determine the metric value corresponding to the branch for which the applied branch

10. A method of generating a branch weight for branches of a trellis for a Viterbi-like detector, wherein the detector is used in a system having Gaussian noise, comprising: selecting a plurality of signal samples, wherein each sample corresponds to a different sampling time instant; calculating a first value representing a logarithm of a quotient of a determinant of a trellis branch dependent covariance matrix of said signal samples and a determinant of a trellis branch dependent covariance matrix of a subset of said signal samples; calculating a second value representing a quadratic of said signal samples less a plurality of target values normalized by a trellis branch dependent covariance of said signal samples; calculating a third value representing a quadratic of a subset of said signal samples less a plurality of channel target values normalized by a trellis branch dependent covariance of said subset of signal samples; calculating the branch weight from said first, second, and third values; and

weight values determined by a trellis branch dependent covariance of said subset of signal samples, calculating the branch weight from said first, second, and third values; and

'839 Patent at Claim 10



- Covariance Matrix  $C_i$  is a "Matrix."

US 6,201,839 B1

7

$$H = \log \det \left( \frac{C_i}{2\pi} \right) + \frac{1}{2} \mathbf{y}_i^T C_i^{-1} \mathbf{y}_i - \frac{1}{2} \mathbf{y}_i^T \mathbf{y}_i$$
 (11)

In the derivations of the branch metrics (9), (10) and (13), an assumption was made on the exact Viterbi-type architecture, that is, the metric can be applied to any Viterbi-type algorithm such as FBSM, FDS/DF, RAM-BSM, or MBSM.

FIG. 3A illustrates a block diagram of a branch metric computation circuit 48 that computes the metric  $M_i$  for a branch of a trellis, as in Equation (13). Each branch of the trellis requires a circuit 48 to compute the metric  $M_i$ .

A logarithmic circuit 50 computes the first term of the right hand side of (13):

$$\left( \log \det \left( \frac{C_i}{2\pi} \right) \right)$$

and a quadratic circuit 52 computes the second term of the right hand side of (13):  $\frac{1}{2} \mathbf{y}_i^T C_i^{-1} \mathbf{y}_i - \frac{1}{2} \mathbf{y}_i^T \mathbf{y}_i$ . The circuit 50 and 52 implement the algorithm of the Viterbi-like detector 30. A summing circuit 54 computes the sum of the outputs of the circuits 50 and 52.

As stated above, the covariance matrix  $C_i$  is given as:

$$C_i = \begin{bmatrix} \sigma_i^2 & \mathbf{w}_i \\ \mathbf{w}_i^T & \mathbf{C}_i \end{bmatrix}$$

Using standard techniques of signal processing, it can be shown that:

$$\frac{\partial \log \det \left( \frac{C_i}{2\pi} \right)}{\partial \sigma_i^2} = -\frac{1}{\sigma_i^2} \mathbf{I}_i$$

This ratio of determinants is referred to as  $\sigma_i^2$ , i.e.:

$$\sigma_i^2 = \frac{\partial \log \det \left( \frac{C_i}{2\pi} \right)}{\partial \sigma_i^2} = -\frac{1}{\sigma_i^2} \mathbf{I}_i$$

It can be shown by using standard techniques of signal processing that the sum of the last two terms of (13), i.e. the output of the circuit 52, is:

$$H = \frac{1}{2} \mathbf{y}_i^T C_i^{-1} \mathbf{y}_i - \frac{1}{2} \mathbf{y}_i^T \mathbf{y}_i$$

Where the vector  $\mathbf{y}_i$  is (1,1)-dimensional and is given by:

$$\mathbf{y}_i = \begin{bmatrix} 1 \\ \mathbf{y}_i \end{bmatrix}$$

Equations (17), (18) and (19) of the circuit 52 can be implemented as a tapped-delay line as illustrated in FIG. 3B. The circuit 52 has L delay elements 54. The tapped-delay line implementation shown in FIGS. 3A and 3B is also referred to as a moving-average, feed-forward, or finite-impulse response filter. The circuit 48 can be implemented using any type of filter as appropriate.

As stated above, the covariance matrix is given as:

$$C_i = \begin{bmatrix} \sigma_i^2 & \mathbf{w}_i \\ \mathbf{w}_i^T & \mathbf{C}_i \end{bmatrix} \quad (14)$$

'839 Patent 7:25-29

the data dependent nature, of the circuit 52. The weights  $\mathbf{w}_i$  and the value  $\sigma_i^2$  can be adapted using three methods. First,  $\mathbf{w}_i$  and  $\sigma_i^2$  can be obtained directly from Equations (20) and (16), respectively, once an estimate of the signal-dependent covariance matrix  $C_i$  is available. Second,  $\mathbf{w}_i$  and  $\sigma_i^2$  can be calculated by performing a Cholesky factorization on the inverse of the covariance matrix  $C_i$ . For example, in the  $L_i D_i^{-1} L_i^T$  Cholesky factorization,  $\mathbf{w}_i$  is the first column of the Cholesky factor  $L_i$  and  $\sigma_i^2$  is the first element of the diagonal matrix  $D_i$ . Third,  $\mathbf{w}_i$  and  $\sigma_i^2$  can be computed directly from the data using a recursive least squares-type algorithm. In the first two methods, an estimate of the covariance matrix is obtained by a recursive least squares algorithm.

'839 Patent 8:10-23

Computing the branch metrics in (10) or (13) requires knowledge of the signal statistics. These statistics are the mean signal values  $m_i$  in (12) as well as the covariance matrices  $C_i$  in (13). In magnetic recording systems, these

'839 Patent 8:24-27

# Noise Covariance Matrices

Claim Term	CMU's Construction	Marvell's Construction
<p>noise covariance matrices</p> <p>'839 Patent Claims 11, 16, 19, 23 '180 Patent Claim 6</p>	<p>noise statistics used to calculate the 'correlation-sensitive branch metrics.'</p> <p>CMU Brf. at 27</p>	<p>covariance matrices of signal samples (where the signal samples include noise).</p> <p>Marvell Brf. at 27-32</p>

- The Dispute
  - ▶ Does "noise covariance matrices" have its ordinary meaning (Marvell) or has it been re-defined in the patent (CMU)?



- First reference to “the noise covariance matrices”
- Implies the ordinary definition

## SUMMARY OF THE INVENTION

⋮

Because the **noise statistics** are non-stationary, the noise sensitive branch metrics are adaptively computed by estimating **the noise covariance matrices** from the read-back data. These covariance matrices are different for each branch of the tree/trellis due to the signal dependent structure of the media noise. Because the channel characteristics in magnetic recording vary from track to track, these matrices are tracked on-the-fly, recursively using past samples and previously made detector decisions.

'839 patent 2:15-23

**1**  
**METHOD AND APPARATUS FOR**  
**CORRELATION-SENSITIVE ADAPTIVE**  
**SEQUENCE DETECTION**  
**CROSS REFERENCE TO RELATED**  
**APPLICATIONS**  
 This application claims priority to Provisional  
 60/440,100, filed May 8, 1997, under 35 U.S.C.  
 119(e).  
**STATEMENT REGARDING FEDERAL**  
**SPONSORED RESEARCH**  
 This invention was supported in part by the  
 Science Foundation under Grant No. IRI-94-0484.  
 United States Government has certain rights in  
 this invention.  
**BACKGROUND OF THE INVENTION**  
 1. Field of the Invention  
 The present invention is directed generally to  
 magnetic recording, sequence detection, and  
 particularly to correlation-sensitive sequence  
 detection.  
 2. Description of the Background  
 In recent years, there has been a major shift in  
 signal detection in magnetic recording. Trellis  
 detection (TD), such as those described in Noll,  
 "A Study of Detection Methods of NRZ Record  
 Data," *IEEE Trans. Magn.*, vol. 18, pp. 1081-108, Jan. 1982,  
 replaced by Viterbi-like detection in the linear  
 response regime. The level of (PML) is achieved  
 between trellis-like detection and decision based  
 (DB) methods such as TD, and TD, and 4.  
 These methods were derived under the assumption  
 of white Gaussian noise (WGN) is present.  
 The resulting detection branch metrics  
 computed in trellis-like detection.  
 It has long been observed that the noise in  
 recording systems is neither white nor stationary.  
 The noise in the media noise results from its  
 signal nature. Combining media noise and its  
 detection has thus far been confined to modifying the  
 branch metrics to account for these effects. 24  
 "Modified Viterbi Algorithm for Inter-Digital  
 Channel," *IEEE Trans. Magn.*, vol. 24, pp. 24-28, pp.  
 Sept. 1982, and Lee et al., "Performance of  
 Modified Trellis Search Algorithm for  
 Detection of Data Replicated Noise," *Proceedings*  
 of the Conference, pp. 563-56, Oct. 1992, and  
 branch metric computation method for combining  
 dependent characteristics of media noise. These  
 methods have been demonstrated on real data.  
 46. "Comparison of Equalization and Channel  
 High-Density Magnetic Recording," *IEEE IS*  
 Conference, New Orleans, April 1997.  
 These methods do not take into consideration  
 the correlation between the samples in the read-back  
 signal. It is well known that the noise in  
 magnetic recording is not white, and its  
 correlation is not zero. This correlation is  
 significant performance degradation in  
 magnetic recording. Thus, there is a need for an  
 adaptive method to track the noise covariance  
 matrices which enables the maximum likelihood sequence detector  
 (MLSD) without making the usual simplifying assumption  
 that the noise samples are independent random variables.

- CMU reads out "covariance matrices" by reading in redundant language ("correlation-sensitive branch metrics")

6. A method of detecting a sequence that exploits a correlation between adjacent signal samples for adaptively detecting a sequence of symbols through a communications channel having intersymbol interference, comprising:

- performing sequence detection on a plurality of signal samples using a plurality of correlation sensitive branch metrics;
- outputting a plurality of correlation sensitive branch metrics;
- outputting a delayed signal sample;
- adaptively updating a plurality of noise covariance matrices in response to the delayed signal samples and the delayed decisions;
- recalculating the plurality of correlation sensitive branch metrics from the noise covariance matrices using subsequent signal samples; and
- repeating steps (a)–(e) for every new signal sample.

statistics used to calculate the correlation-sensitive branch metrics

correlation sensitive branch metrics

See '839 Patent Claims 11, 16, 19, 23; '180 Patent Claim 6

15  
continued

1. A method of detecting a sequence that exploits a correlation between adjacent signal samples for adaptively detecting a sequence of symbols through a communications channel having intersymbol interference, comprising:

(a) performing sequence detection on a plurality of signal samples using a plurality of correlation sensitive branch metrics;

(b) outputting a plurality of correlation sensitive branch metrics;

(c) outputting a delayed signal sample;

(d) adaptively updating a plurality of noise covariance matrices in response to the delayed signal samples and the delayed decisions;

(e) recalculating the plurality of correlation sensitive branch metrics from the noise covariance matrices using subsequent signal samples; and

(f) repeating steps (a)–(e) for every new signal sample.

2. The method of claim 1, wherein the correlation sensitive branch metrics are calculated using a plurality of correlation sensitive branch metrics.

3. The method of claim 1, wherein the correlation sensitive branch metrics are calculated using a plurality of correlation sensitive branch metrics.

4. The method of claim 1, wherein the correlation sensitive branch metrics are calculated using a plurality of correlation sensitive branch metrics.

5. The method of claim 1, wherein the correlation sensitive branch metrics are calculated using a plurality of correlation sensitive branch metrics.

6. The method of claim 1, wherein the correlation sensitive branch metrics are calculated using a plurality of correlation sensitive branch metrics.

7. The method of claim 1, wherein the correlation sensitive branch metrics are calculated using a plurality of correlation sensitive branch metrics.

8. The method of claim 1, wherein the correlation sensitive branch metrics are calculated using a plurality of correlation sensitive branch metrics.

9. The method of claim 1, wherein the correlation sensitive branch metrics are calculated using a plurality of correlation sensitive branch metrics.

10. The method of claim 1, wherein the correlation sensitive branch metrics are calculated using a plurality of correlation sensitive branch metrics.

11. The method of claim 1, wherein the correlation sensitive branch metrics are calculated using a plurality of correlation sensitive branch metrics.

12. The method of claim 1, wherein the correlation sensitive branch metrics are calculated using a plurality of correlation sensitive branch metrics.

13. The method of claim 1, wherein the correlation sensitive branch metrics are calculated using a plurality of correlation sensitive branch metrics.

14. The method of claim 1, wherein the correlation sensitive branch metrics are calculated using a plurality of correlation sensitive branch metrics.

15. The method of claim 1, wherein the correlation sensitive branch metrics are calculated using a plurality of correlation sensitive branch metrics.

# CMU's "i.e." Argument Fails

- Re-defining “noise covariance matrices” as “noise statistics” removes “covariance” and “covariance matrices” from the claims
  - ▶ “[a] claim construction that gives meaning to all the terms of the claim is preferred over one that does not do so.”  
*Merck & Co., Inc. v. Teva Pharm. USA, Inc.*, 395 F.3d 1364, 1372 (Fed. Cir. 2005).
  - ▶ “covariance” and “covariance matrices” have well-known meanings in engineering and statistics
  - ▶ “covariance” and “covariance matrices” are used independently in other claims

See '839 Patent Claim 10.